Chapter 5: Overlooked text types: from fictional texts to real-world discourses[*]

Alon Lischinsky

5.1 Introduction

Behind the enthusiastic adoption of corpus-based approaches in discourse research lies the promise of an ability to explore data more completely and representatively. Traditional methods in discourse studies were primarily designed for delicacy and richness; given the complexity of the links between language use and its social context, and the wide range of linguistic features in which these links are expressed, research tended to focus on the 'detailed analysis of a small number of discourse samples' (Fairclough, 1992: 230). But the depth afforded by such approaches places corresponding limits on breadth of coverage: examining particular excerpts in such detail is only possible at the expense of overlooking everything else that goes on in a given discursive practice. The 'fragmentary [and] exemplificatory' nature of the evidence that can be thus gathered poses considerable problems for generalisation (Fowler, 1996: 8). When texts and features for analysis are selected on the basis of the researcher's intuitive judgement (Marchi & Taylor, 2009: 3), there is no guarantee that they truly represent the distinctive patterns that characterise a discourse (Stubbs, 1997).

Corpus approaches have been instrumental in providing scholars with the tools to go beyond such partial examinations, and obtain reliable evidence of typical patterns of description, evaluation and argumentation across large bodies of text. These advances have been, for the most part, conceptualised in terms of size: using computer-assisted tools allows researchers to identify and retrieve relevant linguistic features in datasets large enough to provide more than a selective characterisation. However, quantity by itself is not a sufficient guarantee of representativeness; however large it may be, a sample will remain partial and incomplete unless it adequately covers the range of genres and contexts in which a given discourse circulates.

Especially for the scholar whose interest lies in the varied ways in which language is used to accomplish particular functions (Partington *et al.*, 2013: 4), making inferences from linguistic evidence to social and cultural practices demands acknowledgement of the diversity and complexity of such practices. In this chapter, I focus on the ways in which corpus-assisted discourse studies (CADS) has sometimes failed to address this complexity 'as seriously as it should' (Leech, 2007: 134). In particular, I argue that overlooking fictional and imaginative genres limits the ability of CADS to explore how individuals are motivated and seduced by the meanings and ideological assumptions of discourse. The following section discusses the difficulties involved in determining representativeness in language data, and illustrates the issues raised by the bias of CADS towards particular genres —typically official, public and factual ones. Section 5.3 discusses how fiction and imagination are central to our understanding of the real world, and sketches some of the complex ways in which readers' affect and attention are engaged by imaginative discourses, while Section 5.4 addresses disciplinary divides about fiction and outlines some of the particular interpretive caveats required to deal with such materials. Finally, Section 5.5 offers an example of how these limitations can be addressed by exploring the role of erotic fiction in the circulation of discourses about gender and sexuality.

---

[*] I would like to thank Juliet Henderson, Elvis C. Gomes and an anonymous reviewer for invaluable feedback.

5.2 The salient and the overlooked in CADS

In the sense I am using the term here, discourse studies is concerned with how language features in the performance of social action, and especially with its role in structuring the conduct of communicative activities and shaping interactions between individuals and groups (cf. Partington *et al.*, 2013: 3). The language employed in a given context is studied as a tangible trace of the ways in which speakers engage with one another —harmoniously or contentiously— for the purpose of coordinating their beliefs and behaviours. CADS in particular seeks to capture the recurring traces left by social routines, ‘the ways in which society creates itself’ (Mahlberg, 2007: 196) by discursively producing and reproducing habitual patterns of understanding and acting. From this point of view, the starting point of the analysis is not linguistic but social (Biber, 1993: 244): what CADS seeks to characterise is not a particular language or linguistic variety, but rather a particular situation, purpose or function repeatedly enacted within a speech community. Assessing the representativeness and completeness of this characterisation therefore requires understanding the complex and messy ways in which texts are linked to the circumstances of their production, circulation and use (Maingueneau, 2010: 150).

It is important to note that, despite the air of mathematical rigour carried by the term, representativeness in corpus linguistics invariably involves messy decisions. In the simplest definition, a sample is representative of a broader population if it shares its characteristics at a smaller scale: for each of the dimensions across which the population varies, the sample should show a distribution similar to that of the whole (Moessner, 2009: 223). However, a precise and principled measurement of this similarity is impracticable in linguistics for two reasons. In the first place, one of the terms of comparison is unmeasurable: the textual universe of a language or linguistic variety as a whole is so large that its actual proportions can never be estimated with certainty (Hunston, 2002: 28). Furthermore, the parameters of variation (from participant demographics to topic, medium, purpose and participation framework) are so many that ensuring that a sample remains representative along all of them would be infeasible (Nelson, 2010: 60). In consequence, attempts to design representative corpora are never accurate in statistical terms. That does not mean, however, that the notion is without value: perfect representativeness may not be attainable, but it can be approximated (Leech, 2007: 140). Even if conceptualised more modestly in terms of balance, it provides a useful regulative ideal for scholars seeking a more comprehensive and less biased image of discursive action.

The choice of variables regarding which representativeness should be prioritised depends, ultimately, on the research question that the evidence intends to answer (Kilgarriff & Grefenstette, 2003: 340–342). In an ambitious proposal for best practices in corpus design, Biber (1993: 245) observes that one particularly relevant factor is how ‘important [a given genre is] in defining a culture’, and his argument seems especially apposite for CADS. While no corpus can fully capture a discursive formation —understood as ‘all the things that are said about a given topic at a given historical period’ (Stubbs, 2001: 165)— it is important to approximate the full range of variation that can be found in this textual universe. Crucially, this entails keeping in view that any specific domain of social life involves many different discursive activities, enacted through a variety of genres in complex assemblages: sequential chains, hierarchies of prestige, repertoires defining specific communities, etc. (Prior, 2009: 17). Ädel (2010) offers the example of political discourse: while executive speeches and parliamentary debates are prototypical exemplars, the means used to make sense of and take positions towards political issues are much broader, from manifestos and pamphlets to media interviews and editorials, bumper stickers and lapel buttons. In a similar manner, discourses of sexuality and gender circulate in a wide range of forms: biology handbooks, reproductive health advice materials, legislation on sexual assault and harassment, dating tips in lifestyle magazines, water-cooler gossip, hallway taunts, etc. Yet only few of these genres are covered in general-purpose corpora, and even custom-built ones are rarely comprehensive. Instead of ‘considering the whole network [of genres] to understand the functioning’ of a specific discourse domain (Maingueneau, 2010: 153), much CADS work is limited to snapshots of particularly salient junctures (Stubbs, 2001: 149).

The temptation to adopt such an approach is understandable, in that it simplifies the interpretive work required to make sense of the evidence. Ensuring that textual data can provide insight into a ‘discursive event as social practice’ (Fairclough, 1995: 134) requires taking into account their context of production and use. In monogeneric corpora, no situational variations complicate interpretation: corpus composition acts as a proxy for the relevant contextual information (Thornbury, 2010: 276). However, snapshots offer only limited possibilities for the comparative analysis that is intrinsic to discourse studies (Partington *et al.*, 2013: 12). Of the three levels that Fairclough identifies for categorising discursive practices —the local context of specific discursive exchanges, the institutional context of a whole organisation or domain, and the wider context of culture—, only the first can be appropriately tackled through the analysis of a single genre; institutions and *a fortiori* cultures can only be captured by exploring broader assemblages.

This exploration can in principle be accomplished incrementally, but CADS has shown a persistent bias towards a restricted set of genres. While committed in theory to a more democratic notion of importance, in practice the majority of work in the field focuses on discourse practices made 'sexy' by their public or official nature (Lee, 2008: 92) such as news reporting, political speech, public and corporate policy, or courtroom discourse. Doubtlessly, there are reasons for this bias: the size of the readership or audience is often a useful proxy for cultural salience, since a text engaged with by an audience of millions —such as mainstream print or broadcast media content— will exert a larger influence than one restricted to a narrow segment of the population (Leech, 2007: 138).[i] Texts intended for widespread consumption may also be particularly useful for CADS because in order to be accessible and understandable to a wide audience they must reproduce —or at least acknowledge — mainstream common-sense assumptions (Baker, 2005: 18). Nonetheless, the disproportionate prevalence of work on such materials leaves open important gaps in our understanding of the way in which discourses circulate in society. Just like traditional approaches in discourse studies were limited by addressing only the highest-profile exemplars, CADS is often partial to the highest-profile genres. This forecloses the possibility of a more dynamic and socially-embedded model of how meanings and attitudes are disseminated, taken up and recontextualised.

An important step towards representativeness would be to reduce the gaps caused by the bias towards the factual and the official. In particular, I would like to argue that fictional genres have rarely been accorded a space commensurate with their cultural salience.

5.3 Fiction, fact and meaning

The corpus linguist seeking to model a particular linguistic variety readily acknowledges fiction as one of the important registers that must be included for a balanced portrayal. The CADS scholar — like the critical discourse analyst more generally (Gupta, 2015: 197)— tends to be less willing; being interested primarily in texts as tangible traces of social action, the relevance of genres that make no claims to actuality seems in principle limited (Maingueneau, 2010: 148). Paradoxically, it is the frequent emphasis on social critique what makes CADS uncritically accept the common-sense principle that 'pre-assigns a low modality' to non-factual texts (Hodge, 1990: 166). But this attitude unfairly marginalises forms of discourse that are essential to the way in which meaning circulates within a society.

Taking his lead from recent work in cultural studies, Richardson (2016) argues that attempts to capture social and political reality must not privilege the actual over the fantastic. The manner in which agents make sense of aspects of the real world —whether nation or anorexia, security or sexual fulfilment— is never built solely on the discourses that (claim to) report the *facts* about it; rather, these factual claims are interlocked in multiple and complex ways with discourses in which imagination plays a central role. Thus, for example, understandings of politics do not draw only on government budgets, population censuses or unemployment figures, but also on narratives that articulate utopian projects of the just (or prosperous, or strong) society that is to be achieved, as well as dystopian visions of the decline and degeneracy that we risk (Glynos *et al.*, 2009: 11-12).

The importance of such fantasies in organising and shaping social action is underscored by Fairclough and Fairclough (2012: 103, emphasis mine), who point out that 'discourses as ways of representing the world do not only describe what social reality *is* but also what it *should* be'. The world, both in its natural and human dimensions, is too complex to be fully apprehended; discursive sense-making reduces the interpretative effort that this complexity requires by selectively drawing attention to specific features and aspects of this world, and especially by defining the situation that the individual occupies and the possibilities for agency that are open within it. Within the interlocking of the actual and the fictional that makes up social life, *imaginaries* thus 'have a central role in the struggle […] for "hearts and minds"' by orienting decision and inspiring action (Sum & Jessop, 2013: 165). The ability of imaginative narratives to 'absorb a reader's full attention, to the point that real-time obligations and concerns are temporarily forgotten' (Toolan, 2009: 195) allows them to engage readers' affect in ways that factual discourses can only rarely achieve.

Imaginaries can grip subjects in two different but connected forms. When explicitly construed not as actual, but rather as possible, imaginative discourses operate as projects or visions: they function as goal premises in processes of practical reasoning (Fairclough & Fairclough, 2012: 107), which in turn recommend specific courses of action. But imaginaries can also gain performative power through institutionalisation and naturalisation; being collectively recognised and embedded in the norms and expectations that govern a given social domain, fantasies gain deontic actuality in that they can effectively constrain or enable specific forms of social agency. A growing body of literature suggests that the distinction between the fictional and the factual is not always reflected in audiences' sense-making: information and evaluations gleaned from fictional media can blend with non-fictional ones in their general knowledge (Marsh *et al.*, 2003).The importance accorded to specific social issues within dramatic plots, for example, can affect audience beliefs about their social salience and significance in the real world even if the fictional nature of these narratives is recognised (Mulligan & Habel, 2013); the ideational dimension of discourse comprehension is largely identical regardless of the modality value assigned to the genre (Jeffries, 2015: 163).

The likelihood of this slippage between the imagined and the actual seems especially great regarding those domains where first-hand knowledge is limited. There are numerous aspects of social life that are rarely open to unmediated encounters, either because of geographical, temporal or social distance, or because they are surrounded by privacy, stigma and taboo. In such cases, it is almost impossible to disentangle the object itself from the skein of narratives and imaginings that preform it in our experience. Phenomena as varied as crime, bereavement, romantic love or sexual passion are not only experienced first, but also more frequently and with greater variety in fiction than in real life; while we have little opportunity to observe them directly and systematically (or perhaps *because* we have little opportunity to observe them), they feature prominently in imaginative discourses such as popular fiction or music lyrics (Edwards, 1994: 242). Attention to such genres, then, can 'stretch critical discourse studies in ways that better reflect the ways that meanings circulate in societies' (Richardson, 2016).

5.4 Literature, style and discourse

Though literature has long played a central role in enquiry into language, it has traditionally been conducted under a separate disciplinary aegis, and the relationship between linguistic and literary research has not always been cordial (Fialho & Zyngier, 2014; Gupta, 2015; Maingueneau, 2010). Nevertheless, the intersection of these interests has received considerable attention since the 1960s under the banner of stylistics.

We can distinguish two ways in which an understanding of literature as discourse has informed this field of research. The first involves adopting the methodological repertoire of discourse studies to address questions of literary criticism, such as the work of Fowler (1989), who employs register analysis to explore characterisation. Such techniques have proved useful to describe various aspects or prose and drama, from speech acts and face threats to the management of dialogic interaction. Though corpus-based approaches remain under-represented in literary stylistics (Fialho & Zyngier, 2014: 331), there is growing interest in these methods, sometimes under other disciplinary labels like digital humanities or computational analysis of style (Biber, 2011: 16; Hoover *et al*., 2014: 3; Toolan, 2009: 4). The focus of analysis in such cases, however, is framed in traditional literary terms, as 'the provision of a basis for fuller understanding, appreciation and interpretation of avowedly literary and author-centred texts' (Carter & Simpson 1989: 6). Questions of social function and impact remain marginal.

A different approach is to adopt literary or (more broadly) fictional materials as data for enquiry into language in use. Explorations of literature as a locus of social action were important in early discourse studies; the same Fowler (1981: 80) encouraged treating 'literature as discourse […] to see the text as mediating relationships between language-users: not only relationships of speech, but also of consciousness, ideology, role and class', and other authors such as van Dijk (1977) or Hodge

(1990) embarked in similar arguments. But engagement with fiction became progressively rarer as the articulation of discourse studies with the social increasingly focused on everyday genres and their common-sense assumptions.

Gupta (2015: 200) examines how this elision of 'the literariness of the social and the socialness of the literary' was related to the contested constitution of discourse studies as separate from literary criticism. Stylisticians justified the social component of their analyses by challenging the idea of *literaturnost*, the distinctive uniqueness of literary language; if the same linguistic features that characterise literature can be found elsewhere as well, there is no principled reason for separate treatment (Fowler 1981: 21; Jeffries & McIntyre, 2010: 2; Simpson, 1993: 2). Critical discourse scholars, on the other hand, founded the relevance of their discipline on a socio-political engagement that excluded the more rarefied and aesthetically-oriented domain of the literary. An aspect downplayed in Gupta's account, however, concerns the particular methodological and epistemological complexities posed by treating fiction as discourse while acknowledging its fictional nature (Talbot, 1995: 28; Sunderland, 2004: 60; Sunderland, 2010: 35). Three important features that make it difficult to draw inferences about real-world discourse practices from fictional materials are the indirect and unusual nature of literary meaning-making; the multiplicity of levels and voices in fiction; and the temporal, local and even ontological dislocation that fiction allows.

The first of these issues is closely connected with the traditional conception of *literaturnost*: literary texts are characterised by the poetic drawing attention to the act of linguistic engagement rather than its function. Through the calculated use of expressions that deviate from conventional linguistic and discursive expectations, literature 'makes form palpable' in order to enhance the enjoyment of perceptual and interpretive activity (Sotirova, 2015: 6). Such foregrounding is, of course, hardly limited to literature, and 'discourse can be norm-breaking in everyday usage for everyday reasons' (Toolan, 2009: 25). But while non-fictional genres typically employ deviation to emphasise certain aspects of the force or sense of the message, literature is often interested challenging expectations about the functional structure of language itself (Cook, 1994: 197). This drawing of attention to the constructed nature of the text makes problematic one of the typical assumptions in discourse work: that the process of ideation is downplayed or naturalised so as to make its results uncritically acceptable to readers (Fairclough & Fairclough, 2012: 121).

At the same time, the comparatively greater singularity of literary texts makes it difficult to treat them, in corpus linguistic fashion, as samples from which generalisable patterns can be drawn (Mahlberg, 2015: 144). Even for analysts that do not share the literary critic's interest in the uniqueness of individual texts, the literary is methodologically challenging simply because it is less predictable and more variable than other forms of text. Nevertheless, it is important to remember that not all fiction is *literary* fiction, and in fact the kind of high literature characterised by conspicuous linguistic foregrounding is the exception rather than the norm both in terms of production and of audience. Not only modern popular genres such as chick lit or detective fiction, but also traditional narrative forms in the oral tradition, follow much stricter (if implicit) rules of composition (Opas & Tweedie, 1999: 89; Semino & Short, 2004: 25). There is no principled reason to assume that the range of variation is so significantly greater than in non-fictional genres that generalisations about fiction are impossible.

Even in forms characterised by predictable narrative formulae there is space for innovation; Leech (1985: 48) conceptualises these as secondary deviations from the reader's expectations about the genre, which informs savvy readers' interpretation and enjoyment of the text (Walsh, 2015: 125). This multiplication of the layers that must be considered to make sense of fiction represents a second challenge for the discourse analyst. At the most obvious level, fiction tends to be polyphonic: rather than consistently expressing a single point of view, it refracts ideas, attitudes and beliefs through a plurality of protagonists and narrative voices that may be in tension or outright conflict with one another (Toolan, 2009: 193–4). In consequence, it is impossible to draw direct links between the presence of a certain representation or propositional content within the discourse and the writer’s commitment to its truth. Even beyond the dialogue explicitly attributed to characters, their perspectives can colour the narrative through a variety of indirect features (Semino & Short, 2004: 10ff), and carefully contextualised analysis may be necessary to identify whose point of view is being represented.

Not only variation between texts, but also within them, becomes then a critical factor in analysis. The first-order meaning created through the actions and utterances of protagonists —the *fabula* in traditional narratological terms— is never conveyed fully or directly; even if presented by an allegedly omniscient narrator, the selection of what is to be told and from whose perspective —the *syuzhet*— represents a second order of meaning that may differ from or even contradict the first through ironic, humorous or satirical presentation. Once again, none of these aspects is categorically exclusive to fiction; multi-voicedness is conspicuous in news discourse, and ironic detachment has long been prominent in advertising. But while CADS work on other genres can hope to smooth out occurrences of these phenomena within a larger body of monologic text, in fiction the refracted form of representation must be taken as the default (Sunderland, 2010: 74).

A final issue when seeking to identify traces of real-world actions and attitudes in fiction is that the latter, by definition, does not deal with the real world. Fiction writers have discretion not only to set their narratives in places and times removed from that of composition —a decision that will colour readers' evaluation and interpretation of the events— but also to choose a setting that differs in important ways from reality (Sunderland, 2010: 51): one in which magic exists, for example, or in which humans are hermaphroditic, or in which the Axis powers won World War 2. The distinction is not, however, absolute. Whether a story is realistic or fantastic, the events and characters it portrays must remain intelligible to readers who will interpret them on the basis of background knowledge drawn from the real world, and the default assumption will be that narrative and reality are congruent unless specific information to the contrary is provided (Tabbert, 2016: 29). Fantasy thus gives the author greater scope for imagination, but the ultimate background for these fabulous elements remains the external reality that readers inhabit (Sunderland, 2004: 61).

Departures from realism are nevertheless of particular interest: like stylistic norm-breaking, deviation ‘on the plane of fiction-building’ (Leech & Short, 2007: 128) highlights aspects of the narrative that will be of significance for its interpretation. A typical way in which such deviations are employed is for allegorical or metaphorical purposes (Stephens, 1992: 248): unrealistic elements in the story must be interpreted as stand-ins for aspects of the real world. Treating fiction as fiction requires then considering the choice of a specific setting —realistic or fantastic— and of the particular generic norms attached to such settings as a potentially significant aspect of the way in which a particular perspective is conveyed.

5.5 Case study: discourses of gender and sexuality in erotic writing

This case study focuses on what the analysis of erotic fiction can contribute to our understanding of discourses about gender and sexuality. As the enormous popular success of E.L. James' *Fifty Shades* trilogy illustrates, such fiction —like other forms of pornography— is an increasingly salient part of contemporary cultural life, where explicit representations of sexual activity have become staples of a range of media forms, from print to photography, film and animation (Attwood, 2011). But while the fictional and unrealistic nature of pornographic narratives is readily apparent to their audiences (McKee, 2010), both scholarly and popular debates about the ‘pornification of society’ have been quick to fixate on the impact they may have on real-world behaviours, attitudes and beliefs about sexuality and gender roles. Critical voices have claimed that pornographic discourses contribute to normalising sexual permissiveness, both in terms of increased interest in sexual matters and of acceptance of casual sex and sexual experimentation (e.g., Zillmann, 2000); a more favourable take is that pornography plays an educational role, serving not only to inform about sexual anatomy and mechanics but also to destigmatise sexual desire and curiosity, especially when other sources of information are lacking, incomplete or perceived as judgemental (Albury, 2014).

5.5.1 Fragmentation and stereotyping in the porn debates

One particularly contentious topic has been the relationship between pornographic discourses and issues of power in sexual relations. Porn has been criticised for ‘eroticising inequality’, denying female sexual agency in line with conventional ideologies, and structuring its representation of women in terms of their attractiveness to men (Crabbe & Corlett, 2010; Gill, 2003). But for all the heat in these debates there has been a surprising scarcity of evidence; more than 25 years after Williams' (1989: 29) complaint that ‘so much has been written about the issue of pornography and so little about its actual texts’, there is still considerable uncertainty about the range of discourses articulated even in its mainstream varieties (for some valuable exceptions, see Baker, 2005; Bolton, 1995; Koller, 2015; Marko, 2008; Morrish & Sauntson, 2007; Motschenbacher, 2010).

In particular, I focus on two questions raised in critiques of gendered representation. The first is that of *fragmentation*. Analyses of sexualisation in media discourse have highlighted how bodies can be dehumanised by dismantling them into disjointed anatomical elements (Caldas-Coulthard, 2008: 465): using body parts as meronymic stand-ins for the whole person dissolves our perception of a unified and conscious subject, so that they appear ‘not as whole people but as fetishized, dismembered “bits”, as objects’ whose volition and humanity are elided (Gill, 2009a: 96). A corpus approach allows us to assess whether such representations are, as is often claimed, especially characteristic of pornography as opposed to other forms of fictional or informative discourse.

The second is that of the discursive construction of *gendered bodies*. Other than those for primary —and, to some extent, secondary— sexual characteristics, terms for body parts do not directly index biological sex; there is nothing in the core semantics of lexemes such as ‘nipple’, ‘thigh’, ‘belly’ or ‘chest’ that limits their reference to either the male or female body. Nevertheless, they may acquire gendered associations on the basis of their typical contexts of appearance (Hellinger & Bußmann, 2011: 11). Recurring practices of reference and description can provide insights on the social or covert gendering that colours stereotypical body talk.

5.5.2 The L1K corpus of erotic fiction

`<INSERT TABLE 5.1 AROUND HERE>`

The data analysed here were collected from *Literotica.com* (2016), one of the oldest, largest and most widely-read erotic fiction repositories online, archiving more than 1.5 million user-contributed stories. Though the writing advice available to contributors ‘represents a normative model of a “good story” as one involving plot and character development, complexity, and non-explicit elements’ (Paasonen, 2010: 144), in practice texts range from elaborate novellas to wall-to-wall sexual accounts. *Literotica* imposes few restrictions on the content it will publish: only bestiality, mutilation, snuff and underage sexual encounters are banned. Within these limits taboo subjects are an ‘object of emotional investment’ (Paasonen, 2011: 109), and some of the most popular categories concern incest, swinging, bondage and sadomasochism. The corpus analysed here comprises the top-rated 39 stories from 26 categories in the archive (excluding texts other than short stories to maintain generic consistency), totalling approximately 10 million word-tokens tagged for part of speech using the NLTK averaged perceptron tagger (Bird *et al.*, 2009);[ii] see Table 5.1 for details.

5.5.3 Fragmented bodies vs whole subjects

`<INSERT FIGURE 5.1 AROUND HERE>`

Fragmentation has been observed in factual genres such as news and advertising (Attenborough, 2011; Caldas-Coulthard, 2008), as well as literature. After illustrating how female bodies are 'dismembered' in both traditional poetry and modern thriller fiction, Mills (1995: 133-5) suggests that this convention is so deeply gendered that it would be 'very difficult to imagine the same process being applied to the depiction of male characters'. From a corpus perspective, this hypothesis can be conceptualised by comparing the frequency of references to characters by a proper name or a personal pronoun (*holonymic* references) with those in which a body part stands in for the whole person (*meronymic* references). Figure 5.1 shows that across the corpus the frequency of each type of reference varies quite widely and largely independent of the other.

`<INSERT TABLE 5.2 AROUND HERE>`

Body part terms in erotic narratives are generally overlexicalised (Marko, 2008), showing the norm-breaking typically associated with literature, but their frequencies follow a typically Zipfian distribution and most occurrences are captured by a relatively limited set of terms. Table 5.2 lists the 50 most frequent ones. Many of these are immediate indicators of the *aboutness* of the corpus —making reference to male and female genitalia, breasts and nipples, buttocks and the anus, and other erogenous zones— though other anatomical terms are frequent as well; as Figure 5.2 shows, both types appear more frequently in *Literotica* stories than in general fiction as represented by the imaginative writing sections of three large reference corpora.

`<INSERT FIGURE 5.2 AROUND HERE>`

At first blush, such evidence seems to support the fragmentation hypothesis: participants in erotic stories are reduced to their parts —especially their private parts— much more frequently than in other forms of writing. Nevertheless, pornographic representations do not only capture their characters in extreme close-ups of bodily action, but also frequently talk about the person as a whole. Holonymic references are in fact more common in L1K than in the reference corpora (Figure 5.3),[iii] and they frequently addresses emotional, volitional and epistemic dimensions that unequivocally involve the characters' subjectivity. Their most frequent right-hand verbal collocates included terms denoting mental (WANT, WONDER, UNDERSTAND, WISH), behavioural (WATCH) and verbal processes (TELL, WHISPER, YELL), together with the more predictable material ones (TAKE, TURN, GO, WALK, WEAR). Even if these actions are less prominent in porn than in other genres (being negative keywords in comparison to the imaginative section of the BNC), they are far from absent. This suggests that the pervasive attention to bodily action and sensation does not come *at the expense* of attention to the subjects' individuality, but appears *in addition* to it. Issues of ‘character motivation, desire, and sexual build-up’ (Paasonen, 2010: 151) provide a sustained counterpoint to the fleshy details of body part talk.

`<INSERT FIGURE 5.3 AROUND HERE>`

In addition, a closer look at meronymic references shows that they are not always or necessarily depersonalising. Kuhn (1985: 36) points out that Western cultural norms recognise the face as ‘stand[ing] in for the person's whole being’, and other forms of partial physical framing have the same effect: eyes are frequently called ‘windows to the soul’, and the heart is often used to refer specifically to the emotional and personal dimensions of subjectivity (Niemeier, 2000). Significantly, these three meronyms occur in *Literotica* stories no less frequently than in other fiction. The narrative gaze does thus not only linger on the naughty bits; its focus on characters' physical presence is often used to provide evidence of their emotions, attitudes and reactions in the form of grinning faces, wide-open eyes or pounding heartbeats. Attention to such aspects is necessary in order to avoid the temptation of a ‘paranoid reading’ (Paasonen, 2011: 134) that simply confirms pre-existing assumptions and criticisms of pornography as dehumanising.

5.5.4 Gendered ideals in body talk

However, this does not mean that such criticisms can be simply ignored. Gill (2009b: 153–4, emphasis mine) argues that ‘claims about the “sexualization of culture” have paid insufficient attention to the *different* ways in which different bodies are represented erotically’. Rather than taking sexualisation as an undifferentiated monolith, analysis must consider how the patterns and modes it adopts intersect with axes of social difference. Though race, class and age are all important in mediating sexualisation, I focus here on the crucial role of gender.

`<INSERT FIGURE 5.4 AROUND HERE>`

Figure 5.4 shows that the hyperbolic carnality of Literotica stories is unequally distributed across this axis. Other than terms for male genitalia, only a few body parts —CHEST, HAND, FINGER, CROTCH— are more characteristic of male representations, and even in such cases the difference is relatively minor, with log ratios ranging from 0.5 to 0.75. Facial features (FACE, EYE, EAR, TONGUE) are mentioned with roughly similar frequency regardless of gender (log ratios between -0.5 and 0.5). All other terms in the list —including references to the lower limbs, buttocks, abdomen, nipples, hair and mouth— are more frequently used to refer to female than to male characters, even if the higher overall frequency of female mentions is factored out. Body talk overall, then, seems to be stereotypically associated with a focus on women; even if most of this vocabulary has a gender-neutral core meaning, the patterns in which it routinely appears associate its semantics with femininity.

<INSERT TABLE 5.3 AROUND HERE>

One way to explore in more detail the nuances that the discursive construction of gender superimposes over biological difference is to focus on terms that are intrinsically gendered by denoting sexually dimorphic aspects: the labelling, description and evaluation of primary and secondary sexual characteristics can illustrate normative ideals of the male and female body. Table 5.3 shows the adjectival collocates most strongly associated with terms for male and female genitalia. Though anatomical differences mean that direct lexical overlaps are unlikely, the most frequent terms on both lists focus on the visible physiological signals of arousal and orgasm. Talk about females shows greater lexical diversity, but sexual anticipation, readiness and pleasure seem to be important in characterisation regardless of gender.

Other aspects such as size, however, show sharp contrasts: male genitalia are hyperbolically large, whereas female genitalia are described as ‘tiny’ or ‘tight’ following the logic of ‘heterosexual structuralism’ that presents male and female bodies in binary opposition (Paasonen, 2011: 125-126). Also noticeable is the importance of the lexical field of grooming in the construction of the feminine sexual ideal: while there is occasional mention of male genital shaving, it is over 40 times more frequent in talk about females. Such representational practices indicate different normative relationships towards the body: while the ideal COCK is born ‘huge’ or ‘massive’, the ideal PUSSY is achieved through extensive investment in womanscaping labour.

<INSERT TABLE 5.4 AROUND HERE>

A second approach involves exploring the purely social gendering discursively applied to parts that are biologically and functionally equivalent in healthy human bodies. The example of EYE, salient in both male and female characterisation, is instructive (Table 5.4). Closing the eyes and opening them widely are frequent indices of emotion without a specific gender association, but other routine formulae are strikingly different. Males' eyes are primarily defined by the direction and manner of their looking: they STARE fixedly at their target, BORE through it, ROAM over it or LINGER on its sexual characteristics. Sometimes this is a result of overpowering attraction, when men can't KEEP their eyes away from the curves of a partner, sometimes deliberate, when they STRAIN their gaze to watch. Discussion of emotional expression —whether gentle or predatory— is much less common. In contrast, women's eyes are more frequently described in terms of appearance or emotion than gaze. Most of their characteristic actions are involuntary: they ROLL in annoyance, FLUTTER in abandon, SPARKLE with excitement or GLAZE with tears, often in the company of other displays of feeling such as parted lips. Emotional distress seems to be a distinctly female condition in the corpus, with collocates related to crying appearing much more frequently in references to women. A binary opposition seems at play here as well: men's eyes are presented as a site of agency and volition, whereas those of women reflect the uncontrollable welling of supervening emotion.

5.6 Evaluation: fictional stories and their real consequences

What then can we learn about the discursive construction of the gendered body from an examination of erotic fiction? The patterns of reference and description found in this corpus are in many cases congruent with those observed in other, more factual and mainstream, discourses. Rather than the misogynistic reduction of women to ‘anonymous, panting playthings, adult toys, dehumanized objects to be used, abused, broken and discarded’ (Brownmiller, 1975: 394) suggested by anti-pornography criticism, the portrayal of female characters in porn is no more extreme in its fragmentation than that of advertising or journalism. The lavish attention devoted to the details of bodily actions and reactions is doubtlessly a significant aspect of these narratives, but interpreting it as a denial of the humanity and subjectivity of characters would be a distorting oversimplification.

The construction of the protagonists involves not just such fleshy details, but also addresses their emotional moods and responses, their cognitive capacities and their communicative engagements. Rather than being elided, as critics have argued, these dimensions appear refracted through the point of view of the narrator and presented primarily through their visible signs; though a systematic examination of the modulation of narrative point of view was beyond the remit of this chapter, corpus methods can reveal the traces of internal focalisation that provide the appearance of first-person witnessing that characterises the genre. Body talk in pornographic narratives is thus overdetermined: a good deal of it functions, in fact, as the main means of conveying the inner life of characters. If anything is neglected in porn, it is not the actions, thoughts and feelings of protagonists, but the details of the background against which they are set. Pornography is, after all, about the graphic representation of sexual activity, and the prominence of body part terms is an obvious reflection of this subject matter.

The need to take into account such contextual factors as the purpose and uses of a genre, of course, not new to discourse analysts, but it comes into a sharper relief when fictional materials are concerned. Far too often, analyses assume that the gendered scripts present in pornographic materials are internalised by their users (e.g., Vannier *et al*., 2014: 254); though it is easy to recognise that erotic narratives are not intended as realistic accounts of actual sexual encounters (Baker, 2005: 154), the temptation remains to interpret them as idealised versions of the sexual activities that authors and readers would like to participate in. However, such interpretations fail to address pornographic fiction as *fiction*. Erotic stories are made tellable —and therefore enjoyable— precisely by the spectacular character of the events and participants they portray, and the often conspicuous lack of realism of its scenarios has to be understood in terms of this specific context of production and use.

From this point of view, the hyperbolic binarism with which porn portrays physical sexual characteristics is more closely linked to its appetite for transgression (Paasonen, 2011) than it is to normative discourses of the properly gendered body; the profusion of massive cocks and tiny pussies is one of the ways in which pornography attempts to make engrossing the ultimately repetitive and predictable dynamics of sexual encounters. Without considering such narrative constraints, ideological interpretations are unreliable. Common-sense interpretations that explain audiences' relationship to porn in terms of identification or ideological reproduction miss the fact that their reactions are often ambivalent, and can involve as much disturbed or confused fascination as outright appeal (Paasonen, 2011: 182). If the study of fiction can stretch our comprehension of the social circulation of discourses, it also requires the analysis to stretch its understanding of how these discourses are used and engaged with, recognising the complexity of attachments that go beyond the tired binary of hegemony and resistance.

Of course, acknowledging the diversity of possible engagements with fiction does not entail denying that the discourses it contains can be naïvely reproduced, but this possibility must be assessed against the background of the other sites and genres in which these discourses circulate. Erotic fiction can certainly function as an ‘instructional discourse’ providing audiences with normative ideas about the characteristics and dispositions they should find desirable in partners— or should adopt themselves in order to be found desirable (Baker, 2005: 190). But such readings are all the more likely because of the absence of alternative spaces where open and non-judgemental discussion of sexual activities can be found. It is the existence of ill-informed —and therefore vulnerable— audiences that underscores the ideological dimension of porn, though it is equally important to bear in mind that even such audiences are equipped with critical literacies developed in their engagement with other genres. Even to relatively naïve readers the hyperbolic excess of pornographic representations may suggest their transgressive and even camp nature. For the critical analyst, perhaps the most problematic aspects of pornography —in terms of reinforcing social norms and expectations— are not the spectacular displays that have attracted critical attention, but rather those features it borrows seamlessly from factual discourses: one of the main contributions that the analysis of fiction can make to CADS is shedding light on the ways in which even our fantasies often remain bound by the assumptions and preconceptions of the society we live in.

References

Ädel, A. 2010. How to use corpus linguistics in the study of political discourse In O'Keeffe, A., & McCarthy, M., eds. *The Routledge Handbook of Corpus Linguistics*. London: Routledge, pp. 591–604.

Albury, K. 2014. Porn and sex education, porn as sex education. *Porn Studies* 1(1–2), 172–181.

Attenborough, F. T. 2011. Complicating the sexualization thesis: The media, gender and 'sci-candy'. *Discourse & Society* 22(6), 659–676.

Attwood, F. 2011. Sex and the media. In Ross, K., ed. *The Handbook of Gender, Sex, and Media*. Oxford: Wiley-Blackwell, pp. 455–469.

Baker, P. 2005. *Public Discourses of Gay Men*. London: Routledge.

Bolton, R. 1995. Sex talk: Bodies and behaviors in gay erotica. In Leap, W. , ed. *Beyond The Lavender Lexicon: Authenticity, Imagination, and Appropriation in Lesbian and Gay Languages*. Luxembourg: Gordon and Breach, pp. 173–206 .

Biber, D. 1993. Representativeness in corpus design. *Literary and Linguistic Computing* 8(4), 243–257.

Biber, D. 2011. Corpus linguistics and the study of literature: Back to the future? *Scientific Study of Literature* 1(1), 15-23.

Bird, S., Klein, E., & Loper, E. 2009. *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*. Sebastopol, CA: O'Reilly.

Caldas-Coulthard, C. R. 2008. Body branded: Multimodal identities in tourism advertising. *Journal of Language and Politics* 7(3), 451–470.

Carter, R., & Simpson, P. 1989. Introduction. In Carter, R., & Simpson, P., eds., *Language, Discourse and Literature: An Introductory Reader in Discourse Stylistics*. London: Unwin Hyman, pp. 1–18.

Cook, G. 1994, *Discourse and Literature.* Oxford: Oxford University Press.

Crabbe, M., & Corlett, D. 2010. Eroticising inequality: Technology, pornography and young people. *Domestic Violence Resource Centre Victoria Quarterly*, 3(Spring), 1–6.

Edwards, E. D. 1994. Does love really stink? The 'mean world' of love and sex in popular music of the 1980s. In Epstein, J., ed. *Adolescents and Their Music: If It's Too Loud, You're Too Old*. New York: Garland, pp. 225–250.

Fairclough, N. 1992. *Discourse and Social Change*. Cambridge: Polity.

Fairclough, N. 1995. *Critical Discourse Analysis: The Critical Study of Language*. London: Routledge.

Fairclough, I., & Fairclough, N. 2012. *Political Discourse Analysis: A Method for Advanced Students*. London: Routledge.

Fialho, O. & Zyngier, S. 2014. Quantitative methodological approaches to stylistics. In Burke, M., ed. *The Routledge Handbook of Stylistics*. London: Routledge, pp. 329–345.

Fowler, R. 1981. L*iterature as Social Discourse: The Practice of Linguistic Criticism*. London: Batsford.

Fowler, R. 1989. Polyphony in *Hard Times*. In Carter, R., & Simpson, P., eds., *Language, Discourse and Literature: An Introductory Reader in Discourse Stylistics*. London: Unwin Hyman, pp. 77–94.

Fowler, R. 1996 On critical linguistics. In Caldas-Coulthard, R., & Coulthard, M., eds., *Texts and Practices: Readings in Critical Discourse Analysis*. London: Routledge, pp. 3–14.

Gill, R. 2003. From sexual objectification to sexual subjectification: The resexualisation of women's bodies in the media. *Feminist Media Studies* 3(1), 100–106.

Gill, R. 2009a. Beyond the 'sexualization of culture' thesis: An intersectional analysis of 'sixpacks', 'midriffs' and 'hot lesbians' in advertising. *Sexualities* 12(2), 137–160.

Gill, R. 2009b. Supersexualize me! Advertising and the mid-riffs. In Attwood, F., ed. *The Sexualization of Western Culture: Mainstreaming Sex*. London: I.B. Taurus, pp. 93–111.

Glynos, J., Howarth, D., Norval, A., & Speed, E. 2009. *Discourse Analysis: Varieties and Methods*. ESRC National Centre for Research Methods review paper 14. Southampton: National Centre for Research Methods. Retrieved from http://eprints.ncrm.ac.uk/796/

Gupta, S. 2015. *Philology and Global English Studies: Retracings*. Basingstoke: Palgrave Macmillan.

Hellinger, M., & Bußmann, H. 2001. The linguistic representation of women and men. In Hellinger, M., & Bußmann, H., eds., *Gender Across Languages: The Linguistic Representation of Women and Men*. Amsterdam & Philadelphia, PA: John Benjamins, pp. 1–26.

Hodge, R. 1990. *Literature as Discourse*. Cambridge: Polity Press.

Hoover, D. L., Culpeper, J., & O'Halloran, K. 2014. *Digital Literary Studies: Corpus Approaches to Poetry, Prose, and Drama*. London: Routledge.

Hunston, S. 2002. *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.

Jeffries, L., & McIntyre, D. 2010. *Stylistics*. Cambridge: Cambridge University Press.

Jeffries, L. 2015. Critical stylistics. In Sotirova, V., ed.. *The Bloomsbury Companion to Stylistics*. London & New York: Bloomsbury, pp. 157–176.

Kilgarriff, A., & Grefenstette, G. 2003. Introduction to the special issue on the web as corpus. *Computational Linguistics* 29(3), 333–347.

Koller, V. 2015. The subversive potential of queer pornography. A systemic-functional analysis of a written online text. *Journal of Language and Sexuality* 4(2), 254–271.

Kuhn, A. 1985. *The Power of the Image: Essays on Representation and Sexuality*. London: Routledge & Kegan Paul.

Lee, D. Y. W. 2008. Corpora and discourse analysis: new ways of doing old things. In Bhatia, V., Flowerdew, J., & Jones, R. H., eds. 2008. *Advances in Discourse Studies*. London: Routledge, pp. 86–99 .

Leech, G. N. 1985. Stylistics. In van Dijk, T. A., ed. *Discourse and Literature*. Amsterdam & Philadelphia, PA: John Benjamins, pp. 39–57.

Leech, G. N. 2007. New resources, or just better old ones? The Holy Grail of representativeness. In Hundt, M., Nesselhauf, N., & Biewer, C., eds. *Corpus Linguistics and the Web*. Amsterdam & New York: Rodopi, pp. 133–149.

Leech, G. N., & Short, M. 2007. *Style in Fiction: A Linguistic Introduction to English Fictional Prose*. 2nd ed. Harlow: Pearson Education.

Mahlberg, M. 2007. Lexical items in discourse: identifying local textual functions of sustainable development. In Hoey, M., Mahlberg, M., Stubbs, M., & Teubert, W., eds., *Text, Discourse and Corpora: Theory and Analysis*. London: Continuum, pp. 191–218.

Mahlberg, M. 2015. Corpus stylistics. In Sotirova, V., ed. *The Bloomsbury Companion to Stylistics*. London & New York: Bloomsbury, pp. 139–156.

Maingueneau, D. 2010. Literature and discourse analysis. *Acta Linguistica Hafniensia* 42(S1), 147–158.

Marchi, A., & Taylor, C. 2009. If on a winter's night two researchers…: A challenge to assumptions of soundness of interpretation. *Critical Approaches to Discourse Analysis across Disciplines* 3(1), 1–20.

Marko, G. 2008. *Penetrating Language: A Critical Discourse Analysis of Pornography*. Tübingen: Narr.

Marsh, E. J., Meade, M. L., & Roediger III, H. L. 2003. Learning facts from fiction. *Journal of*

Memory and Language, 49(4), 519–536.

McKee, A. 2010. Does pornography harm young people? *Australian Journal of Communication* 37(1), 17–36.

Mills, S. 1995. *Feminist stylistics*. London: Routledge.

Moessner, L. 2009. How representative are the *Philosophical Transactions of the Royal Society* of 17th–century scientific writing? In Renouf, A., & Kehoe, A., eds. *Corpus Linguistics: Refinements and Reassessments*. Amsterdam: Rodopi, pp. 221–238.

Morrish, E., & Sauntson, H. 2007. *New Perspectives on Language and Sexual Identity*. Basingstoke & New York: Palgrave Macmillan.

Motschenbacher, H. 2010. *Language, Gender and Sexual Identity: Poststructuralist Perspectives*. Amsterdam & Philadelphia, PA: John Benjamins.

Mulligan, K., & Habel, P. 2013. The implications of fictional media for political beliefs. *American Politics Research* 41(1), 122–146.

Nelson, M. 2010. Building a written corpus. In O'Keeffe, A., & McCarthy, M., eds. *The Routledge Handbook of Corpus Linguistics*. Abingdon & New York: Routledge, pp. 53–65.

Niemeier, S. 2000. Straight from the heart: metonymic and metaphorical explorations. In Barcelona, A., ed. *Metaphor and Metonymy at the Crossroads. A Cognitive Perspective*. Berlin: Mouton de Gruyter, pp. 195–211.

Opas, L. L., & Tweedie, F. 1999. The magic carpet ride: reader involvement in romantic fiction. *Literary and Linguistic Computing* 14(1), 89–101.

Paasonen, S. 2010. Good amateurs: erotica writing and notions of quality. In Attwood, F., ed. *Porn.com: Making Sense of Online Pornography*. New York: Peter Lang, pp. 138–154.

Paasonen, S. 2011. *Carnal Resonance. Affect and Online Pornography*. Cambridge, MA & London: The MIT Press.

Partington, A., Duguid, A., & Taylor, C. 2013. *Patterns and meanings in discourse: theory and practice in corpus-assisted discourse studies (CADS)*. Amsterdam & Philadelphia, PA: John Benjamins.

Prior, P. 2009. From speech genres to mediated multimodal genre systems: Bakhtin, Voloshinov, and the question of writing. In Bazerman, C., Bonini, A., & Figueiredo, D. *Genre in a changing world*. WAC Clearinghouse & Parlor Press: Fort Collins, CO & West Lafayette, IA, pp.17-34.

Richardson, J. E. 2016. Culture, British fascism and Critical Discourse Studies. Paper presented at the CADAAD 2016 conference. Catania: Università degli Studi di Catania.

Semino, E., & Short, M. 2004. *Corpus Stylistics: Speech, Writing and Thought Presentation in a Corpus of English Writing*. London: Routledge.

Simpson, P. 1993. *Language, Ideology, and Point of View*. London: Routledge.

Sotirova, V. 2015. Introduction. In Sotirova, V., ed. *The Bloomsbury Companion to Stylistics*. London & New York: Bloomsbury, pp. 3–18.

Stephens, J. 1992. *Language and Ideology in Children's Fiction*. London & New York: Longman.

Stubbs, M. 1997. Whorf's children: Critical comments on Critical Discourse Analysis (CDA). In Ryan, A., & Wray, A., eds. *Evolving models of language: Papers from the Annual meeting of the British association for applied linguistic held at the University of Wales, Swansea, September 1996*. Clevedon: BAAL, pp. 110–116.

Stubbs, M. 2001. *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford: Blackwell.

Sum, N.-L., & Jessop, B. 2013. *Towards a Cultural Political Economy. Putting Culture in its Place in Political Economy*. Cheltenham & Northampton, MA: Edward Elgar.

Sunderland, J. 2004. Gendered discourses in children’s literature. *Gender Studies* 1(3), 60–84.

Sunderland, J. 2010. *Language, Gender and Children's Fiction*. London & New York: Continuum.

Tabbert, U. 2016. *Crime and Corpus: The Linguistic Representation of Crime in the Press*. Amsterdam & Philadelphia, PA: John Benjamins.

Talbot, M. 1995. *Fictions at Work: Language and Social Practice in Fiction*. London: Longman.

Thornbury, S. 2010. What can a corpus tell us about discourse? In O'Keeffe, A., & McCarthy, M., eds. *The Routledge Handbook of Corpus Linguistics*. London: Routledge, pp. 270–287.

Toolan, M. J. 2009. *Narrative Progression in the Short Story: A Corpus Stylistic Approach*. Amsterdam & Philadelphia, PA: John Benjamins Publishing.

van Dijk, T. 1977. The pragmatics of literary communication. Forastieri-Braschi, E., Guinness, G., & López-Morales, H., eds. *On Text and Context*: *Papers of the International Conference on Literary Communication*. Río Piedras: Editorial Universitaria, pp. 3–16.

Vannier, S. A., Currie, A. B., & O’Sullivan, L. F. 2014. Schoolgirls and soccer moms: a content analysis of free ‘teen’ and ‘MILF’ online pornography. *The Journal of Sex Research* 51(3), 253–264.

Walsh, C. 2015. Feminist stylistics. In Sotirova, V., ed. *The Bloomsbury Companion to Stylistics*. London & New York: Bloomsbury, pp. 122–138.

Williams, L. 1989. *Hard Core: Power, Pleasure, and the Frenzy of the Visible.* Berkeley, CA: University of California Press.

Zillmann, D. 2000. Influence of unrestrained access to erotica on adolescents' and young adults' dispositions toward sexuality. *Journal of Adolescent Health* 27*,* 41–44.

[i] Texts that have a restricted circulation may still have broader indirect effects if they are taken up and reproduced by other, wider-reaching voices; political manifestos, for example, are read far less often than the media reports on their content. Nevertheless, this uptake should leave its own tangible traces in a well-designed corpus.

[ii] The author is grateful to Mark Allen Thornton, Princeton University, for scraping the materials and metadata.

[iii] For ease of visualisation I present only pronominal data in the following analyses, but the inclusion of nominal references does not introduce any significant differences.